VARIATIONAL REPRESENTATIONS
OF VARADHAN FUNCTIONALS

HAROLD BELL AND WLODZIMIERZ BRYC

(Communicated by Claudia Neuhauser)

Abstract. Motivated by the theory of large deviations, we introduce a class
of non-negative non-linear functionals that have a variational “rate function”
representation.

1. Introduction

Let \((X, d)\) be a Polish space with metric \(d()\) and let \(C_b(X)\) denote the space of
all bounded continuous functions \(F : X \to \mathbb{R}\). In his work on large deviations of
probability measures \(\mu_n\), Varadhan [12] introduced a class of non-linear functionals
\(L\) defined by

\[
L(F) = \lim_{n \to \infty} \frac{1}{n} \log \int_X \exp(nF(x))d\mu_n
\]

and used the large deviations principle of \(\mu_n\) to prove the variational representation

\[
L(F) = L_0 + \sup_{x \in X} \{ F(x) - I(x) \},
\]

where \(I : X \to [0, \infty]\) is the rate function governing the large deviations, and
\(L_0 := L(0) = 0\).

Several authors [1, 3, 4, 9, 10, 11] abstracted non-probabilistic components from
the theory of large deviations. In particular, in [3] (see also [10, Theorem 3.1]) we
give conditions which imply the rate function representation (2) when the limit (1)
exists, and we show that the rate function is determined from the dual formula

\[
I(x) = L(0) + \sup_{F \in C_b(X)} \{ F(x) - L(F) \}.
\]

In fact, one can reverse Varadhan’s approach, and show that large deviations of
probability measures \(\mu_n\) follow from the variational representation (2) for (1) (see
[8, Theorem 1.2.3]). In this context we have \(\mu_n(X) = 1\) which implies \(L(0) = 0\) in
(1) and correspondingly \(L_0 = 0\) in (2).

“Asymptotic values” in [3] are essentially what we call Varadhan Functionals
here; the theorems in that paper are not entirely satisfying because the assumptions
are in terms of the underlying probability measures. In this paper we present a more
satisfying approach which relies on the theory of probability for motivation purposes only.

**Definition 1.1.** A function \( L : \mathcal{C}_b(X) \to \mathbb{R} \) is a Varadhan Functional if the following conditions are satisfied:

(4) \( \text{If } F \leq G, \text{ then } L(F) \leq L(G) \) for all \( F, G \in \mathcal{C}_b(X) \),

(5) \( L(F + \text{const}) = L(F) + \text{const} \) for all \( F \in \mathcal{C}_b(X) \), \( \text{const} \in \mathbb{R} \).

Expression (4) provides an example of Varadhan Functional, if the limit exists. Another example is given by variational representation (2).

Condition (4) is equivalent to \( L(F \_ G) = L(F) \_ L(G) \), where \( a \_ b \) denotes the maximum of two numbers. Varadhan Functionals like (1) satisfy a stronger condition.

**Definition 1.2.** A Varadhan Functional \( L \) is maximal if the following condition is satisfied:

(6) \( L(F \_ G) = L(F) \_ L(G) \).

It is easy to see that each Varadhan Functional \( L(\cdot) \) satisfies the Lipschitz condition \( |L(F) - L(G)| \leq \|F - G\|_\infty \); compare (3). Thus \( L \) is a continuous mapping from the Banach space \( \mathcal{C}_b(X) \) of all bounded continuous functions into the real line. We will need the following stronger continuity assumption, motivated by the definition of the countable additivity of measures.

**Definition 1.3.** A Varadhan Functional is \( \sigma \)-continuous if the following condition is satisfied:

(7) \( \text{If } F_n \searrow 0, \text{ then } L(F_n) \to L(0) \).

Notice that if \( X \) is compact, then by Dini’s theorem and the Lipschitz property, all Varadhan Functionals are \( \sigma \)-continuous.

Maximal Varadhan Functionals are convex; this follows from the proof of Theorem 2.1 which shows that formula (2) holds true for all Varadhan Functionals when the supremum is extended to all \( x \) in the Čech-Stone compactification of \( X \).

A simple example of convex and maximal but not \( \sigma \)-continuous Varadhan Functional is \( L(F) = \limsup_{x \to -\infty} F(x) \), where \( F \in \mathcal{C}_b(\mathbb{R}) \). This Varadhan Functional cannot be represented by variational formula (2). Indeed, (2) implies that \( \mathbb{I}(x) \geq F(x) - L(F) = F(x) \) for all \( F \in \mathcal{C}_b(\mathbb{R}) \) that vanish at \( \infty \); hence \( \mathbb{I}(x) = \infty \) for all \( x \in \mathbb{R} \) and (2) gives \( L(F) = -\infty \) for all \( F \in \mathcal{C}_b(\mathbb{R}) \), a contradiction.

An example of a convex and \( \sigma \)-continuous but not maximal Varadhan Functional is \( L(F) = \log \int_X \exp F(x) \nu(dx) \), where \( \nu \) is a finite non-negative measure.

### 2. Variational representations

The main result of this paper is the following.

**Theorem 2.1.** If a maximal Varadhan Functional \( L : \mathcal{C}_b(X) \to \mathbb{R} \) is \( \sigma \)-continuous, then there is \( L_0 \in \mathbb{R} \) such that variational representation (2) holds true and the rate function \( \mathbb{I} : X \to [0, \infty] \) is given by the dual formula (3). Furthermore, \( \mathbb{I}(\cdot) \) is a tight rate function: sets \( \mathbb{I}^{-1}([0, a]) \subset X \) are compact for all \( a > 0 \).

The next result is closely related to Bryc [3, Theorem T.1.1] and Deuschel & Stroock [6, Theorem 5.1.6]. Denote by \( \mathcal{P}(X) \) the metric space (with Prokhorov
metric) of all probability measures on a Polish space \( X \) with the Borel \( \sigma \)-field generated by all open sets.

**Theorem 2.2.** If a convex Varadhan Functional \( \mathbb{L} : \mathcal{C}_b(X) \to \mathbb{R} \) is \( \sigma \)-continuous, then there is a lower semicontinuous function \( \mathcal{J} : \mathcal{P}(X) \to [0, \infty] \) and a constant \( L_0 \) such that

\[
\mathbb{L}(F) = L_0 + \sup_{\mu \in \mathcal{P}} \left\{ \int F \, d\mu - \mathcal{J}(\mu) \right\}
\]

for all bounded continuous functions \( F \).

A well-known example in large deviations is the convex \( \sigma \)-continuous functional \( \mathbb{L}(F) := \log \int \exp F(x) \nu(dx) \) with the rate function in (8) given by the relative entropy functional

\[
\mathcal{J}(\mu) = \begin{cases} 
\int \log \frac{d\mu}{\nu} \, d\mu & \text{if } \mu \ll \nu \text{ is absolutely continuous}, \\
\infty & \text{otherwise}.
\end{cases}
\]

**Remark 2.1.** Deuschel & Stroock [6, Section 5.1] consider convex functionals : \( \mathcal{C}_b(X) \to \mathbb{R} \) such that (const) = const. Such functionals satisfy condition (5). Indeed, write \( F + const \) as a convex combination \( F + const = (1 - \theta)F + \frac{\theta}{2} (2 const) + \frac{\theta}{2} (2F) \), where \( 0 < \theta < 1 \). Using convexity and \( \Phi(const) = const \) we get \( \Phi(F + const) \leq \Phi(F) + const + \theta (\frac{\Phi(2F)}{\theta} - \Phi(F)) \). Since \( \theta > 0 \) is arbitrary this proves that \( \Phi(F + const) \leq \Phi(F) + const \). By routine symmetry considerations (replacing \( F \mapsto F - const \), and then const \( \mapsto -const \)), (5) follows.

### 3. Proofs

Let \( L_0 := \mathbb{L}(0) \). Passing to \( \mathbb{L}'(F) := \mathbb{L}(F) - L_0 \) if necessary, without losing generality, we assume \( \mathbb{L}(0) = 0 \).

**Lemma 3.1.** Let \( \hat{X} \) be a compact Hausdorff space. Suppose \( X \subset \hat{X} \) is a separable metric space in the relative topology. If \( x_0 \in \hat{X} \setminus X \), then there are bounded continuous functions \( F_n : \hat{X} \to \mathbb{R} \) such that

- (i) \( F_n(x) \searrow 0 \) for all \( x \in X \),
- (ii) \( F_n(x_0) = 1 \) for all \( n \in \mathbb{N} \).

**Proof.** Since \( \hat{X} \) is Hausdorff, for every \( x \in X \) there is an open set \( U_x \ni x \) such that its closure \( \bar{U}_x \) does not contain \( x_0 \).

By Lindelöf property for separable metric space \( X \), there is a countable subcover \( \{U_n\} \) of \( \{U_x\} \).

A compact Hausdorff space \( \hat{X} \) is normal. So there are continuous functions \( \phi_n : \hat{X} \to \mathbb{R} \) such that \( \phi_n \big|_{\bar{U}_n} = 0 \) and \( \phi_n(x_0) = 1 \).

To end the proof take \( F_n(x) = \min_{1 \leq k \leq n} \phi_k(x) \).

The following lemma is contained implicitly in [6, Theorem T.1.2].

**Lemma 3.2.** Theorem 2.1 holds true for compact \( X \).
Proof. Let \( \mathbb{I}(\cdot) \) be defined by (3). Thus \( \mathbb{I}(x) \geq F(x) - \mathbb{L}(F) \) which implies \( \mathbb{L}(F) \leq \sup_{x \in X} \{ F(x) - I(x) \} \). To end the proof we need therefore to establish the converse inequality. Fix a bounded continuous function \( F \in C_b(X) \) and \( \epsilon > 0 \). Let \( s = \sup_{x \in X} \{ F(x) - I(x) \} \). Clearly \( F(x) - I(x) \leq s \leq \mathbb{L}(F) \). By (3) again, for every \( x \in X \), there is \( F_x \in C_b(X) \) such that \( I(x) < F_x(x) - \mathbb{L}(F_x) + \epsilon \). Therefore

\[
F(x) \leq s + I(x) < s + \epsilon + F_x(x) - \mathbb{L}(F_x).
\]

This means that the sets \( U_x = \{ y \in X : F(y) - F_x(y) < s + \epsilon - \mathbb{L}(F_x) \} \) form an open covering of \( X \). Using compactness of \( X \), there is a countable number of open balls of radius \( \delta < \epsilon \) for all \( F \). Let \( F_1 \in \mathbb{L}(F) \) be a bounded continuous function such that \( \mathbb{L}(F_1) = \mathbb{I}(x) \) for all \( x \in X \). Then, writing \( F_i = F_{x(i)} \) we have

\[
F(x) < \max_{1 \leq i \leq k} \{ F_i(x) - \mathbb{L}(F_i) \} + s + \epsilon
\]

for all \( x \in X \).

Using (1), (2), and (3) we have

\[
\mathbb{L}(F) \leq \mathbb{L} \left( \max_{1 \leq i \leq k} \{ F_i - \mathbb{L}(F_i) \} + s + \epsilon \right)
= \mathbb{L} \left( \max_i \{ F_i - \mathbb{L}(F_i) \} \right) + s + \epsilon
= \max_i \{ \mathbb{L}(F_i) \} + s + \epsilon.
\]

Since (3) implies \( \mathbb{L}(F_i - \mathbb{L}(F_i)) = \mathbb{L}(F_i) - \mathbb{L}(F_i) = 0 \) this shows that \( s \leq \mathbb{L}(F) < s + \epsilon \). Therefore \( \mathbb{L}(F) = s \), proving (2).

Proof of Theorem 2.1. Let \( \hat{X} \) be the Čech-Stone compactification of \( X \). Since the inclusion \( X \subset \hat{X} \) is continuous, we define \( \hat{\mathbb{I}} : C_b(\hat{X}) \rightarrow \mathbb{R} \) by \( \hat{\mathbb{I}}(\hat{F}) := \mathbb{I}(\hat{F}|X) \).

It is clear that \( \hat{\mathbb{I}} \) is a maximal Varadhan Functional, so by Lemma 3.2 there is \( \hat{\mathbb{I}} : \hat{X} \rightarrow [0, \infty) \) such that \( \hat{\mathbb{I}}(\hat{F}) = \sup_{x \in \hat{X}} \{ \hat{F}(x) - \hat{I}(x) : x \in \hat{X} \} \).

Using \( \sigma \)-continuity (7) it is easy to check that \( \hat{\mathbb{I}}(x) = \infty \) for all \( x \in \hat{X} \setminus X \). Indeed, given \( x_0 \in \hat{X} \setminus X \) by Lemma 3.1 there are \( F_n \in C_b(X) \) such that \( F_n \not\uparrow 0 \) on \( X \), but \( F_n(x_0) = C > 0 \). Then from (3) we get \( \hat{\mathbb{I}}(x_0) \geq \mathbb{L}(0) + F_n(x_0) - \mathbb{L}(F_n) \rightarrow \mathbb{L}(0) + C \).

Since \( C > 0 \) is arbitrary, \( \hat{\mathbb{I}}(0) = \infty \).

This shows that \( \hat{\mathbb{I}}(\hat{F}) = \sup_{x \in \hat{X}} \{ \hat{F}(x) - \hat{I}(x) : x \in \hat{X} \} \) for all \( \hat{F} \in C_b(\hat{X}) \). It remains to observe that since \( \hat{X} \) is a Čech-Stone compactification, every function \( F \in C_b(X) \) is a restriction to \( X \) of some \( \hat{F} \in C_b(\hat{X}) \) (see [7 IV.6.22]). Therefore (2) holds true for all \( F \in C_b(X) \).

To prove that the rate function is tight, suppose that there is \( a > 0 \) such that \( \mathbb{I}^{-1}[0,a] \) is not compact. Then there is \( \delta > 0 \) and a sequence \( x_m \in X \) such that \( \mathbb{I}(x_m) \leq a \), and \( d(x_m, x_n) > \delta \) for all \( m \neq n \). Since Polish spaces have Lindelöf property, there is a countable number of open balls of radius \( \delta/2 \) which cover \( X \). For \( k = 1, 2, \ldots \), denote by \( B_k \supseteq x_k \) one of the balls that contain \( x_k \), and let \( \phi_k \) be a bounded continuous function such that \( \phi_k(x_k) = 2a \) and \( \phi_k = 0 \) on the complement of \( B_k \). Then \( F_n = \max_{k \geq n} \phi_k \not\uparrow 0 \) pointwise. On the other hand (2) implies \( \mathbb{L}(F_n) \geq L_0 + F_n(x_n) - I(x_n) \geq L_0 + a \), contradicting (7).

Lemma 3.3. If \( \mathbb{I}(\cdot) \) is a Varadhan Functional, then

\[
\inf_{x \in X} \{ F(x) - G(x) \} \leq \mathbb{L}(F) - \mathbb{L}(G).
\]
Proof. Let $\text{const} = \inf_x \{ F(x) - G(x) \}$. Clearly, $F \geq G + \text{const}$. By positivity condition \[(\text{3})\] this implies $L(F) \geq L(G + \text{const}) = L(G) + \text{const}$. \hfill \qed

The next lemma is implicitly contained in the proof of \[3\, \text{Theorem T.1.1}\]. Let $\mathcal{P}_a(X)$ denote all regular finitely-additive probability measures on $X$ with the Borel field.

Lemma 3.4. If $L(\cdot)$ is a convex Varadhan Functional on $C_b(X)$, then there exist a lower semicontinuous function $\mathcal{J} : \mathcal{P}_a(X) \to [0, \infty]$ such that

\[(10) \quad \mathcal{J}(F) = L(0) + \sup\{ \mu(F) - \mathcal{J}(\mu) : \mu \in \mathcal{P}_a(X) \},\]

and the supremum is attained.

Proof. Let $\mathcal{J}(\cdot)$ be defined by

\[(11) \quad \mathcal{J}(\mu) = L(0) + \sup\{ \mu(F) - L(F) : F \in C_b(X) \}\]

and fix $F_0 \in C_b(X)$. Recall that throughout this proof we assume $L(0) = 0$.

By the definition of $\mathcal{J}(\cdot)$, we need to show that

\[(12) \quad L(F_0) = \sup_{\mu} \inf_L \{ \mu(F_0) - \mu(F) + L(F) \},\]

where the supremum is taken over all $\mu \in \mathcal{P}_a(X)$ and the infimum is taken over all $F \in C_b(X)$. Moreover, since \[(\text{11})\] implies that $\mathcal{J}(\mu) \geq \mu(F_0) - L(F_0)$ for all $\mu \in \mathcal{P}_a(X)$, then $L(F_0) \geq \sup_\mu \inf_F \{ \mu(F_0) - \mu(F) + L(F) \}$. Hence to prove \[(\text{12})\], it remains to show that there is $\nu \in \mathcal{P}_a(X)$ such that

\[(13) \quad L(F_0) \leq \nu(F_0) - \nu(F) + L(F) \quad \text{for all} \quad F \in C_b(X)\]

(also, for this $\nu$, the supremum in \[(\text{11})\] will be attained). To find $\nu$, consider the following sets. Let

$\mathcal{M} = \{ F \in C_b(X) : \inf_x [F(x) - F_0(x)] > 0 \}$

and let $\mathcal{N}$ be a set of all finite convex combinations of functions $g(x)$ of the form $g(x) = F(x) + L(F_0) - L(F)$, where $F \in C_b(X)$.

It is easily seen from the definitions that $\mathcal{M}$ and $\mathcal{N}$ are convex; also $\mathcal{M} \subset C_b(X)$ is non-empty since $1 + F_0 \in \mathcal{M}$, and open since $\{ F : \inf_x [F(x) - F_0(x)] \leq 0 \} \subset C_b(X)$ is closed. Furthermore, $\mathcal{M}$ and $\mathcal{N}$ are disjoint. Indeed, take arbitrary $g \in \mathcal{N}$ and let $\alpha_k$ be the convex combination.

Then

$$\inf_{x} \{ g(x) - F_0(x) \}$$

$$= \inf_{x} \left\{ \sum \alpha_k F_k(x) - F_0(x) \right\} - \sum \alpha_k L(F_k) + L(F_0)$$

$$\leq \inf_{x} \left\{ \sum \alpha_k F_k(x) - F_0(x) \right\} - \sum \alpha_k L(F_k) + L(F_0) \leq 0,$$

where the first inequality follows from the convexity of $L(\cdot)$ and the second one follows from \[(\text{9})\] applied to $F = \sum \alpha_k F_k(x)$ and $G = F_0$. 

\[\]
Therefore \( M \) and \( N \) can be separated, i.e. there is a non-zero linear functional \( f^* \in C_b^*(X) \) such that for some \( \alpha \in \mathbb{R} \)

\[
f^*(N) \leq \alpha < f^*(M)
\]

(14) (see e.g. [7 V. 2. 8]).

**Claim:** \( f^* \) is non-negative.

Indeed, it is easily seen that \( F_0(\cdot) \) belongs to \( N \), and, as a limit of \( \epsilon + F_0(x) \) as \( \epsilon \to 0 \), \( F_0 \) is also in the closure of \( M \). Therefore by (14) we have \( \alpha = f^*(F_0) \). To end the proof take arbitrary \( F \) with \( \inf_x F(x) > 0 \). Then \( F+L(F_0) \in M \) and by (14)

\[
f^*(F) = f^*(F+F_0) - f^*(F_0) > \alpha - f^*(F_0) = 0.
\]

This ends the proof of the claim.

Without losing generality, we may assume \( f^*(1) = 1 \); then it is well known (see e.g. [2 Ch. 4 Section 4 Theorem 1]) that \( f^*(F) = \nu(F) \) for some \( \nu \in \mathcal{P}_b(X) \); for regularity of \( \nu \) consult [7 IV.6.2].

**Proof of Theorem 2.2.** Lemma 3.4 gives the variational representation (10) with the supremum taken over a too large set. To end the proof we will show that \( J(\mu) = \infty \) on measures \( \mu \) that fail to be countably-additive.

Suppose that \( \mu \) is additive but not countably additive. Then Daniell-Stone theorem implies that there is \( \delta > 0 \) and a sequence \( F_n \not\in 0 \) of bounded continuous functions such that \( \int F_n d\mu > \delta > 0 \) for all \( n \). By (11) and \( \sigma \)-continuity, \( J(\mu) \geq \mathbb{L}(0) + C \int F_n d\mu - \mathbb{L}(CF_n) \geq \mathbb{L}(0) + \frac{\delta}{C} - \mathbb{L}(CF_n) \to \mathbb{L}(0) + C\delta \). Since \( C \) is arbitrary, then \( J(\mu) = \infty \) for all \( \mu \) that are additive but not countably-additive. Thus (10) implies (8).

**References**


DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CINCINNATI, P.O. BOX 210025, CINCINNATI, OHIO 45221–0025
E-mail address: bellh@email.uc.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CINCINNATI, P.O. BOX 210025, CINCINNATI, OHIO 45221–0025
E-mail address: bryewz@email.uc.edu